Preference for Flexibility from Incomplete Resolution of Uncertainty

Min-Feng Lu   Hendrik Rommeswinkel*

October 24, 2022

Abstract

We analyze information and flexibility preferences in a Savage framework. Preference for flexibility is commonly modeled with subjective states that may be completely unrelated to the objective state space. We instead adopt a framework in which preference for flexibility is due to incomplete resolution of uncertainty from the objective state space before making a choice from a menu. We characterize both state-dependent and state-independent representations of (indirect) expected utility maximization. In both cases, the decision relevant hidden information that the decision maker expects to obtain prior to a choice from the menu is uniquely identified.
Keywords: Knowledge, Flexibility, Uncertainty, Information

1 Introduction

In many settings, decision makers prefer to gain more information to make better decisions. In some cases, which information a decision maker uses can be of legal or policy relevance. For example, the use of race and gender information by managers may motivate stricter antidiscrimination laws. The use of inside information by an investor may be criminal. However, for analysts it is often difficult to identify what information a decision

*Both authors: Department of Economics and Center for Research in Econometric Theory and Applications, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Taipei 106.

maker uses to make choices. In this paper, we show that under very general conditions a decision maker's preferences for *further* information can inform an analyst about the information the decision maker expects to gain.

Consider the example of a scientist from the Massachusetts Institute of Technology who was convicted of inside trading (**newsreport**). According to several news reports, the scientist was found out by their google search history which included searches for how the SEC investigates inside trading. Thus, his *information seeking* behavior revealed his possession of decision relevant information that was hidden for the analysts. In our model we formalize how all decision-relevant hidden information of a decision maker can be identified by their preferences over binary menus combined with information partitions.

Decisions are often made sequentially and preceding choices determine what the agent can choose afterward and what information is available at the next decision stage. Rational decision makers will generally want to have more information and more flexibility in future stages. Commonly, these different aspects are kept separate in the decision theory literature. The literature on preference for flexibility infers from menu preferences a potential information structure that a decision maker uses to optimally choose from menus. The literature on information preferences following instead analyzes when information preferences can be induced by a subsequent decision problem. The present paper brings together these two strands of the decision theory literature; the preference over information following (Blackwell, 1953) and the literature on preference for flexibility following (Kreps, 1979).

We employ a Savage (1954) model of decisions under uncertainty but enrich this framework by also eliciting menu preferences and information preferences. In our model, an agent faces a two stage decision problem. In the second stage, the agent knows that the true state of the world is within some event and chooses a standard Savage act from a menu. In the first stage, the decision maker chooses an information partition and a menu conditional on every information set that may occur. Thus, a first stage act may be preferred to another first stage act because it offers better information, more flexibility in the menus on some events, or the subsequent acts in a menu lead to better outcomes on some events.

We characterize three decision models. In our first model, we assume that the indirect utility property holds for the menu preferences but weaken Savage's axioms to allow for the value of outcomes to be state-dependent. The decision maker's preferences can then be represented by a sum across events of the information partition of the value obtained from choosing the optimal subsequent act based on a state-dependent expected utility function. Naturally, beliefs are not uniquely identified in this case.

In our second model, we weaken the indirect utility property by restricting it to singleton menus but use analogous versions of Savage's monotonicity and likelihood outcome independence axioms to guarantee a state-independent utility over outcomes. We further impose that information is only instrumentally valuable. The decision maker's preferences can then be represented by expected indirect utility over subsequent acts with a unique probability measure.

In our third model, we allow for an intermediate case between the usual preference for flexibility models with a completely subjective state space and our second model. We allow for additional information being retrieved by the agent between the first and second stage without an analyst observing this information retrieval. We therefore call this information "hidden". In this model, we relax the indirect utility axiom and we impose that both flexibility and information have a nonnegative value and that the decision maker's information preferences fulfill a consistency requirement. Our axioms allow for the value of information to be zero when it would be instrumentally useful. We show that the decision maker's preferences can be represented by expected indirect utility maximization but that before making a choice from a menu, the decision maker uses a uniquely characterized information sigma algebra to refine the information partition. Thus, instead of a completely subjective state space our decision maker's state space when making a final menu choice is embedded in the objective state space and from preferences it can be determined whether an agent expects to be able to distinguish two objective states before making a final decision.

Lastly, we discuss for our state-dependent model whether in practice it is possible to uniquely identify the hidden information of a decision maker. Menu preferences are well known to pose difficulties in practical applications because the number of questions an individual has to answer for an analyst to retrieve their menu preferences explodes as the number of options increases. Moreover, state dependent models are well known to pose difficulties when it comes to unique identification of utilities and these problems could in principle extend to the identification of hidden information. We show that this is not the case and that indeed all decision relevant hidden information is uniquely identifiable from preferences over first stage acts that yield only binary menus.

Gilboa and Lehrer (1991), Liang (2019), and Rommeswinkel et al. (2020) discuss the existence of subsequent decision problems justifying the information preferences of an expected utility maximizing decision maker. Kreps (1979), Dekel et al. (2001), and Nehring (1999) discuss under which conditions on menu preferences there exists information that an expected utility maximizing decision maker expects to receive prior to the menu choice.[1] In this paper, we discuss under which conditions on preferences over menus and information there exists a hidden information partition that an expected utility maximizing decision maker uses to choose from menus. Differently put, we analyze when the subjective state space of preference for flexibility models can be embedded into the objective state space and when the subsequent decision problem that induces information preferences is a choice from the (objective) menus. Dillenberger et al. (2014) use menu preferences to elicit the information an expected utility maximizing decision maker anticipates to receive before making a decision from menus. Their state space is finite and acts map into (objective) utilities while we employ an infinite state space and utilities over outcomes are subjective. Our elicitation method however requires the analyst to offer different information partitions. Expected utility maximization with hidden information is also discussed by van Zandt (1996) and Morris (1997).

The paper proceeds as follows. We introduce our notation in section 2 and our decision models in section 3. Our axioms are discussed in section 4 and are used in section 5 to characterize the decision models. Section 7 shows that given any finite set of subsequent acts we can elicit all decision relevant hidden information even in the presence of state dependent utilities.

2 Notation

The decision maker faces uncertainty about the state of the world. Let $\mathcal{S}$ be the set of possible *states* of the world and $\mathcal{E}$ be an atomless σ-algebra of *events* on $\mathcal{S}$.

A partition $\mathcal{I}$ of $\mathcal{S}$ is called an *information partition* if for all $I \in \mathcal{I}$, $I \in \mathcal{E}$.

[1] Closest to the present paper are Nehring (1999) and Rommeswinkel et al. (2020) which study preference for flexibility and preference for knowledge, respectively, in a Savage framework.

$\mathcal{I}$ is a *refinement* of $\mathcal{J}$, denoted $\mathcal{I} \leq \mathcal{J}$ if for all $I \in \mathcal{I}$, there exists $J \in \mathcal{J}$ such that $I \subseteq J$.

Let $\mathcal{F}$ be a set of *subsequent acts*. In the main part of the paper, we assume no further structure on each subsequent act $f, g, \ldots$. In the appendix that treats the state-independent case, we assume that subsequent acts are standard Savage acts mapping states into outcomes. A *menu* m of subsequent acts is a finite subset of $\mathcal{F}$. Denote by $\mathcal{M}$ the set of all menus $m, n, \ldots$.

The decision maker has preferences over an information partition combined with possibly different menus at every event in the information partition. For this, we define an *information act* as a mapping $a : \mathcal{I} \to \mathcal{M}$ where $\mathcal{I}$ is a finite information partition. For an information act a, the information partition $\iota(a)$ is the domain of a. The definition of an information act therefore differs from the usual decision-theoretic act by its domain being an information partition instead of the state space. Denote the set of all information acts $a, b, \ldots$ by $\mathcal{A}$. With a slight abuse of notation, we denote by $a^{-1}(m) = \bigcup\{E \in \iota(a) : a(E) = m\}$ the event on which m is offered to the decision maker. An information act with information partition $\{E^1, E^2, \ldots, E^p\}$ and corresponding menus $m^1, \ldots, m^p$ is denoted by $m^1_{E^1} m^2_{E^2} \cdots m^p_{E^p}$. As standard in the literature, the last subscript (E^p in the above case) is omitted. As a consequence, a menu m also denotes an information act $m_{\mathcal{S}}$. If G is an event, $a = m^1_{E^1} m^2_{E^2} \cdots m^p_{E^p}$ and $b = n^1_{F^1} n^2_{F^2} \cdots n^q_{F^q}$ are information acts, then $a_G b$ denotes the information acts $m^1_{E^1 \cap G} \cdots m^p_{E^p \cap G} n^1_{F^1 \cap G^c} \cdots n^q_{F^q \cap G^c}$. Thus, $a_G b$ is the information act with an information partition in which states in G are distinguishable from $\overline{G}$, states within G are distinguishable according to the information partition $\iota(a)$ and menus are assigned according to a.

It is noteworthy that the decision maker can never infer any additional information from the available options in the menu about what the state of the world is. Since information acts map elements of an information partition into menus, the information set generated by the preimages of menus $a^{-1}(m)$ is coarser than $\iota(a)$.

An event E is *null* if for all f, g, $f_E g \sim g$. If μ is a measure on $\mathcal{E}$, then for any event E with $\mu(E) > 0$, $m|E$ is the *conditional measure* such that $\mu|E(F) = \frac{\mu(F \cap E)}{\mu(E)}$ for all F.

A set function $v : \mathcal{E} \to \mathbb{R}$ is monotonely continuous if for all sequences (E^k), $v(\bigcup_{k=1}^{l} E^k) \to v(\bigcup_{k=1}^{\infty} E^k)$ as $l \to \infty$.

3 Decision Models

A natural starting point is an additive representation where for a menu m available at information set E, the decision maker gains a utility $v(m, E)$.

Definition 1 (Event Dependent Utility)**.** $\succsim$ on the set $\mathcal{A}$ defined on events $\mathcal{E}$ has an event dependent utility representation if there exist a function $U : \mathcal{A} \to \mathbb{R}$ unique up to affine transformations, a function $v : \mathcal{M} \times \mathcal{E} \to \mathbb{R}$ monotonely continuous in its second argument such that $U(a) \geq U(a')$ if and only if $a \succsim a'$ and

$$U(a) = \sum_{E \in \iota(a)} v\big(a(E), E\big). \tag{1}$$

A decision maker who uses the information supplied by the information act in order to choose optimally from m has a utility representation:

Definition 2 (Indirect Utility)**.** $\succsim$ on the set $\mathcal{A}$ defined on events $\mathcal{E}$ has an indirect utility representation if there exist a function $U : \mathcal{A} \to \mathbb{R}$ unique up to affine transformations, a function $v : \mathcal{F} \times \mathcal{E} \to \mathbb{R}$ monotonely continuous in its second argument, such that $U(a) \geq U(a')$ if and only if $a \succsim a'$ and

$$U(a) = \sum_{E \in \iota(a)} \max_{f \in m} v\big(f, E\big). \tag{2}$$

For our main decision model, we require some additional definitions. An information sigma algebra $\mathcal{H}$ is a sigma algebra on $\mathcal{S}$ that is a coarsening of $\mathcal{E}$, i.e., all elements of $\mathcal{H}$ are also elements of $\mathcal{E}$. The idea of an information sigma algebra is that two elements s, s' of $\mathcal{S}$ are distinguishable in an information sigma algebra $\mathcal{H}$ if there exist disjoint events E and F in $\mathcal{H}$ containing s and s', respectively. Our motivation for introducing information sigma algebras is that we can integrate over its events instead of using summation across elements of an information partition. Summing across information partitions would require that the knowledge gained about events does not contain null sets. For example, if a decision maker expects to gain information about a random variable on the real line and expects to know the exact state if the random variable is nonnegative and gain no information otherwise, then the information partition would be $\{(-\infty, 0)\} \cup \mathbb{R}_+$. Clearly, in this case we would run into difficulties using summation across elements of the information partition. Capturing the information a decision maker expects to gain in a sigma algebra directly provides us with a method of aggregating across events even when the corresponding information partition would contain null sets.

Let $\mathcal{H}$ be an information sigma algebra and E an event. Denote by $\mathcal{H}|E$ the conditional information sigma algebra containing the intersections of $\mathcal{H}$ with E. Let $\mu : \mathcal{E} \to [0,1]$ be a measure and $\mu(E) > 0$, then $\mu^{\mathcal{H}|E} : \mathcal{H}|E \to [0,1]$ fulfills for all $H \in \mathcal{H}$: $\mu^{\mathcal{H}|E}(E \cap H) = \frac{\mu(E \cap H)}{\mu(E)}$.

Definition 3 (State Dependent Hidden Information Utility)**.** $\succsim$ on the set $\mathcal{A}$ defined on events $\mathcal{E}$ has a state dependent hidden information utility representation if there exist a function $U : \mathcal{A} \to \mathbb{R}$ unique up to affine transformations, a function $u : \mathcal{F} \times \mathcal{S} \to \mathbb{R}$ integrable in its second argument, and a unique hidden information sigma algebra $\mathcal{H}$ such that $U(a) \geq U(a')$ if and only if $a \succsim a'$ and

$$U(a) = \sum_{E \in \iota(a)} \int_E \max_{f \in a(E)} w_{f_E} \, \mathrm{d}\mu^{\mathcal{H}|E}. \tag{3}$$

where $w_{f_E}(H) = \int_H u(f,s) \, \mathrm{d}\mu^{\mathcal{E}|H}$.

In principle, it is also possible to consider the case in which a decision maker has hidden outside options. However, in this case we generally lose many uniqueness properties of the utility representations because hidden subsequent decision problems of knowledge preferences are generally not unique. We consider representations with both hidden information and hidden actions an interesting field for further research.

4 Axioms

We assume the usual weak order axiom to guarantee that the decision maker's preference is well behaved.

Axiom 1 (Weak Order)**.** $\succsim$ is complete and transitive.

We impose the following variation of Savage (1954)'s sure-thing principle.

Axiom 2 (Sure-Thing Principle)**.** For all $E \in \mathcal{E}$ and all a, b, c, $d \in \mathcal{A}$,

$$c_E a \succsim c_E b \quad \Longleftrightarrow \quad d_E a \succsim d_E b. \tag{4}$$

It is noteworthy that in all of the involved acts the decision maker will be able to distinguish whether event E obtains or not. This is because $c_E a$ is an information act in which the information partition consists of the information partition of c after intersection with E and the information partition of a after intersection with E^c.

The Sure-Thing Principle allows us to define a relation $\succsim_E$ conditional on an event E: $a \succsim_E b$ if and only if $a_E c \succsim b_E c$ for some c. Such preferences have the interpretation of what the preferences over acts would be in case the individual is informed that the true state of the world is within E. These conditional preferences extend naturally to menus as we can define $m \succsim_E m'$ if and only if the conditional information acts a and a' yielding m and m', respectively, on event E fulfill $a \succsim_E a'$. Similarly, for subsequent acts $f, g \in \mathcal{F}$, we can define $f \succsim_E g$ if and only if $\{f\} \succsim_E \{g\}$.

We focus on a model in which the subsequent acts are valued state-dependently. In the appendix, we also provide a version in which the value of outcomes is state-independent and the structure of subsequent acts is known to the analyst.

We strengthen Savage's Nontriviality axiom by requiring that subsequent acts cannot be indifferent both on a nonnull event and all of its nonnull subevents.

Axiom 3 (Strong Nontriviality)**.** For all $f, g \in \mathcal{F}$ and all nonnull events $E \in \mathcal{E}$, if $f \sim_E g$, then for some nonnull event $F \subset E$, $f \succ_F g$.

We naturally do not assume Savage's Monotonicity and Likelihood Outcome Independence axioms. Appendix **??** provides these axioms and characterizes a state-independent representation.

To introduce the continuity properties we assume, we next define what it means for sequences of events, subsequent acts, menus, and information acts to converge.

Definition 4 (Convergence of Events)**.** A sequence (E^k) of events converges in subjective belief to E, denoted $E^k \rightsquigarrow E$, if the following two conditions hold:

(E1) If $F \subsetneq E$ is nonnull, then there exists N such that for all $k > N$, $F \cap E^k \neq \emptyset$.

(E2) If F is nonnull and for each N, there exists $k > N$ such that $F \subsetneq E^k$, then $F \cap E \neq \emptyset$.

The first condition requires that in the limit, the set E should not be too "large". On the other hand, the second condition prevents the set E from being too "small". Note that our defintion allows a sequence of events converge to infinitely many different events, but the difference between each of them is a null event in terms of the agent's subjective belief. An example of convergence in subjective belief are monotonically increasing or monotonically decreasing sequences of sets of states.[2]

Definition 5 (Convergence of Information Partitions)**.** A sequence (I^k) of information partitions converges in subjective belief to I, denoted $I^k \rightsquigarrow I$, if for each $E_i \in I$, there exists a sequence (E_i^k) such that:

1. $E_i^k \in I^k \cup \{\emptyset\}$ for all k, and
2. $E_i^k \rightsquigarrow E_i$.

Definition 6 (Convergence of Information Acts)**.** A sequence (a^k) of information acts converges in subjective belief to a, denoted $a^k \rightsquigarrow a$, if the following two conditions hold:

1. $\iota(a^k) \rightsquigarrow \iota(a)$.
2. $(a^k)^{-1}(f) \rightsquigarrow a^{-1}(f)$ for all $f \in \mathcal{F}$. where (E_i^k) is the sequence that converges to E_i.

Equipped with these definitions, we can now state our continuity axiom:

Axiom 4 (Continuity)**.** If $a^k \rightsquigarrow a$, $b^k \rightsquigarrow b$, and for all k, $a^k \succsim b^k$, then $a \succsim b$.

Naturally, a decision maker has a nonnegative value of information when using information to solve a decision problem. We therefore assume:

Axiom 5 (Positive Value of Information)**.** For every menu m and events E, F,

$$m_E m_F \succsim_{E \cup F} m_{E \cup F}$$

Similarly, additional options are in the worst case of no value and in the best case provide additional value. Therefore we assume:

Axiom 6 (Positive Value of Flexibility)**.** For all menus m, m' and every event E, if $m \subseteq m'$, then $m' \succsim_E m$.

[2]See appendix G.

If the obtained knowledge is of no use for the agent to make her choice in the second stage, then we assume the value of such knowledge is zero. In particular, this occurs when the menu is singleton. The following axiom captures this idea:

Axiom 7 (Instrumental Knowledge Property)**.** For all events $E, F \in \mathcal{E}$ with $E \cap F = \emptyset$,

$$\{f\}_{E\cup F} \sim_{E\cup F} \{f\}_E\{g\}_F.$$

This effectively allows us to ignore the special structure of the information acts whenever the information act only yields singleton menus.

In case the decision maker receives no hidden information before making the choice from the menu, the decision maker does not care about flexibility unless it provides a new optimal action to take. This can be expressed via the following condition.

Definition 7 (Indirect Utility Property)**.** For every event E, every subsequent act f and every menu m,

$$\{b\} \succsim_E m \Longrightarrow \{b\} \cup m \sim_E \{b\}.$$

We weaken this condition to allow for decision makers who receive hidden information before choosing from the menu.

Axiom 8 (Hidden Indirect Utility Property)**.** Let m and n be two menus and E an event. If $m \cup n \succ_E m$ and $m \cup n \succ_E n$, then there exists an event I such that

1. $m \cup n \sim_{E\cap I} m$,
2. $m \cup n \sim_{E\cap I^c} n$,
3. $m \cup n \sim_E m_I n$, and
4. for all $o \in \mathcal{M}$ and all $F \in \mathcal{E}$, $o_{I\cap F}o \sim_F o$.

The first two conditions are reminiscent of the usual indirect utility property; for some subevent $E \cap I$ the decision maker always chooses from m and for its relative complement the decision maker chooses from n. The third condition requires that the decision maker has no gain of value from being informed about I. The fourth condition requires that the decision maker never gains any value from being informed about I, irrespective of the menu offered and other information provided.

We call the sets I fulfilling the four conditions of Axiom 8 for menus m and n and any event E the hidden information sets identified by menu $m \cup n$ and denote their collection by $\mathcal{H}_{m\cup n}$. We denote by $\mathcal{H}_0$ the set of all hidden identified sets.

5 Axiomatization

In this section, we present our three main representation theorems. To start out, we characterize an event-dependent evaluation of menus and information partitions which provides a starting point for our characterizations. The representation is additively separable in the events of the information partition that the act provides. The value of each additive component may contain any of the following: the likelihood the decision maker attaches to an event, the value of the options given that the event is known, and the intrinsic value of knowing that the event obtains. In such a model, it is impossible to disentangle these aspects of the utility of the decision maker.

Theorem 1. *Suppose $\succsim$ satisfies 9. Then the following statements are equivalent:*

1. *$\succsim$ satisfies Axioms 1, 2, and 4*
2. *$\succsim$ has an event dependent utility representation unique up to affine transformations.*

This representation serves as a starting point for our remaining characterizations. The main difficulty in obtaining this result are the relatively weak continuity assumptions.

If we in addition assume that the indirect utility property holds, then one can show that on every information set the decision maker evaluates a menu by the value of its best alternative.

Proposition 1. *Let $\succsim$ fulfill Axiom 9. Then the following statements are equivalent:*

1. *$\succsim$ satisfies Axioms 1, 2, 4, and Definition 7.*
2. *$\succsim$ has an event utility representation.*

Third, we characterize a representation in which the decision maker expects to gain some information other than $\iota(a)$ provided by the information act before making the final decision. We continue to assume that outcomes are valued state-independently and that the decision maker values knowledge about events only instrumentally. However, we weaken the indirect utility property to allow for knowledge about some events to have no value because the decision maker already expects to gain this knowledge anyway.

Theorem 2. *Let $\succsim$ fulfill Axiom 9. Then the following statements are equivalent:*

1. *$\succsim$ satisfies Axioms 1,2, 4 and 5-8.*
2. *$\succsim$ has a hidden information utility representation.*

The proof roughly proceeds as follows. We first obtain an event dependent utility representation from Theorem 1. The additive components in turn are additively across hidden information sets. Thus, if we were to provide additional information that the decision maker anyways expects to gain, there is no increase in value and thus the additive components are superadditive across events. Together with the positive value of knowledge axiom, requiring the value of knowledge to be subadditive, it follows then that for hidden information sets the additive components are additive. Since the structure of hidden information sets is a sigma algebra, we obtain a representation of the value of a menu in the form of an integral. We finally show that the representation of the value of a menu takes the form of an expected utility of subsequent acts.

6 Comparative Statics

We can compare decision makers by their preference for information, their preference for flexibility.

Definition 8 (Comparative Preference for Information)**.** $\succsim_1$ exhibits more preference for information than $\succsim_2$ if for all E and all α, β, $m = \{\alpha_E\beta, \beta_E\alpha\}$ $m \sim_1 m_E m$ implies $m \sim_2 m_E m$.

Thus, decision maker 1 has a stronger preference for information than decision maker 2 if every hidden information set for decision maker 1 is also a hidden information set for decision maker 2.

Note that stronger definitions will frequently yield an empty comparison relation. For example, suppose we assume that for an arbitrary menu the decision maker 1 strictly prefers to refine an information partition whenever decision maker 2 strictly prefers to refine an information partition. It is straightforward to show that even if both decision makers have the same expected utility preferences over singleton menus and decision maker 2's hidden information refines that of decision maker 1, there are menus where the combination of decision maker 2's hidden information with a partition $\mathcal{I}$ is beneficial but the combination of decision maker 1's hidden information with $\mathcal{I}$ is of no value.

Definition 9 (Comparative Preference for Flexibility)**.** $\succsim_1$ exhibits more preference for flexibility than $\succsim_2$ if for all b and all m it holds that $m_{\mathcal{S}} \succsim_2 \{b\}_{\mathcal{S}} \Rightarrow m_{\mathcal{S}} \succsim_1 \{b\}_{\mathcal{S}}$.

The definition requires that if decision maker 2 prefers a menu to an act, then also decision maker 1 prefers the menu. Generally, two decision maker will only be comparable in their preference for flexibility as long as they have identical preferences over acts. For some menus, an option may only be valuable if one has some information. If decision maker 1 has this information but decision maker 2 does not, then for a properly chosen act, decision maker 1 will prefer the menu to the act but decision maker 2 does not.

Theorem 3. *Let $\succsim_1$ and $\succsim_2$ have hidden information utility representations. Then the following statements are equivalent.*

1. *u_1 is an affine transformation of u_2 and $\mu_1 = \mu_2$. The hidden information partitions H_1 and H_2 fulfill $H_1 \geq H_2$.*

2. *$\succsim_2$ exhibits more preference for information than $\succsim_1$ and both decision makers have identical preferences over singleton acts.*

3. *$\succsim_1$ exhibits more preference for flexibility than $\succsim_2$.*

4. *For some $\gamma, \beta \in X$, $\gamma \succ_1 \beta$, $\gamma \succ_2 \beta$ and for all m, $D(E,m) = \frac{v_2(E,m)}{v_2(\mathcal{S},\gamma)-v_2(\mathcal{S},\beta)}$ is subadditive.*

Proof. 1 $\Rightarrow$ 2: Since both decision makers have up to an affine transformation identical expected utility representations of preferences over singletons, their preferences on singleton acts coincide. Since $H_1 \geq H_2$, every hidden information set of decision maker 2 is also a hidden information set of decision maker 1.

2 $\Rightarrow$ 3: Since both decision makers have identical preferences over singleton acts, it is straightforward to show that their preferences over

subsequent acts coincide conditional on every event E. For simplicity, denote the expected utility preferences over subsequent acts conditional on an event F as $EU_1(\cdot|F) = EU_2(\cdot|F)$. Suppose now that decision maker 2 prefers a menu to an act but decision maker 1 does not. Since decision maker 1's information partition refines that of decision maker 2, for every $H \in \mathcal{H}_2 \cap E$, $\max_{b \in m(H)} EU_2(b|H) \leq \sum_{H' \in \mathcal{H}_1 : H' \subseteq H} \max_{b \in m(H)} EU_1(b|H')$, a contradiction.

$3 \Rightarrow 1$: If the expected utility representations do not coincide up to an affine transformation, then for two subsequent acts such that $b \succ_1 b'$ and $b' \succ_1 b$, we can by continuity find an intermediate act b'' such that $b \succ_1 b'' \succ_1 b'$ and $b' \succ_2 b'' \succ_2 b$, yielding a contradiction to decision maker 1 having more preference for flexibility for the trivial menu $\{b''\}$. Suppose now that $H_1 \not\geq H_2$. Then for some $H \in H_2$, H_1 does not contain H. The menu $\{\gamma_H \beta, \beta_H \gamma\}$ now offers a higher utility to decision maker 2 than to decision maker 1 and we can find an act such that decision maker 2 prefers the menu to the act but decision maker 1 does not. $\square$

We can see that comparative preference for flexibility automatically implies that expected utility preferences are identical. This is not the case for comparative information preferences. Thus, in practice it may be often more desirable to identify hidden information using our definition of hidden information sets.

We can also compare preference for information given a fixed menu. Suppose the value of information is higher for DM1 than for DM2 for all information sets. This value can be measured by equivalent improvements in a comparable act. A consistent value requires the EU preferences of both DMs to be identical. Clearly, DM1 cannot have more decision relevant information than DM2. Thus, for every two states s and s' if decision maker 1 can distinguish between these two states, then there is no value for decision maker 2 to distinguish between these two states. Thus, the best action must already be chosen on these states by DM 2 without any further information. Moreover, every information that is valuable to DM2 must also be valuable to DM1. When we inform DM2 about event E and DM2 chooses a better option on E, then DM1 also chooses a better option on E and gains at least as much in doing so.

7 Elicitation of Hidden Information

In many economic applications, the subsequent acts are not objectively given. For example, if the subsequent choice consists of stocks, then we do not know how the value of the stock relates to each event affecting the stock. Even if we were to incorporate the price of the stock into the state space, we would need to ask the individual many hypothetical questions to properly identify beliefs and utilities. It turns out that allowing for the value of the stock to be state dependent still permits partly eliciting the hidden information that the decision maker expects to receive prior to choosing from the menu.

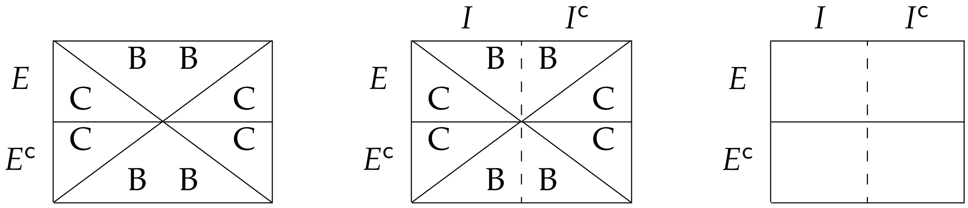
We assume throughout this section that preferences over outcomes are asymmetric in every state which greatly simplifies the exposition. To identify the hidden information, we propose an approach that uses a binary menu $m = \{b^1, b^2\}$. First, from the decision-maker's preferences over information acts, we can obtain a partition $\{E^1, E^2\}$ such that $b^1 \succ b^2$ in and only in E^1. We then pick $F^1 \subseteq E^1$ and $F^2 \subseteq E^2$, and let $F = F^1 \cup F^2$. In the absence of the hidden information, we would have

$$v(m, F) < v(m, F^1) + v(m, F^2)$$

because it is valuable for the decision maker to know whether to choose b^1 or b^2. However, if the decision-maker knows F^1 and F^2 due to her hidden information, then we instead have equality for the above inequality. We therefore gradually change F^1 and F^2 to recover the information sets which obtain equality in the expression above, a process which we describe in more detail below. Finally, repeating the above steps with all binary menus we are able to recover a hidden information partition $\mathcal{P}_2$ based on binary menus.

Our approach has its limitations. If the decision-maker's hidden information G is a subset of some $E \in \mathcal{P}_2$, then this approach will fail to identify G. This may be the case if there is no pair of subsequent acts for which it would be useful to know whether G obtains. We conjecture that there is no approach to identify such G. On the other hand, if we can successfully identify a hidden information set G, then G must intersect some $E \in \mathcal{P}_2$ and therefore we will call G a border-crossing information set if for some acts b, b' it intersects both the event on which b and on which b' is optimal.

We can recover a complete binary relation over subsequent acts in each state.[3] Since we assume transitivity and a strict order of singletons, the binary relation in each state is a linear order and hence has a unique maximal element (if exists). States with the same weak order have the same maximal element while states with the same maximal element may not have the same weak order.



The left table is what the analyst knows given the decision-maker's report. E and E^c are the information that the analyst can offer through an information act. The region marked by B is the collection of states in which the decision-maker prefers the subsequent act b over c and vice versa. The right table is what the decision-maker has in mind. E and E^c are the information given by the analyst while I and I^c are the hidden information.

If the utility of $(E : \{b, c\}, E^c : \{b\})$ is higher than both $(E : \{b\}, E^c : \{b\})$ and $(E : \{c\}, E^c : \{b\})$, then the analyst can infer that there exists hidden information. That is, the decision-maker is able to choose different subsequent acts due to her hidden information. Without loss of generality, assume the decision-maker chooses b in $E \cap I$ and c in $E \cap I^c$, as illustrated below.

	I	I^c
E	b	c
E^c	b	b

Now we would like to identify the region B in the upper right block, i.e., $B \cap E \cap I^c$. Our method is as follows. We pick a small subset $\varepsilon \subseteq B \cap E$ and offer another information act $(E - \varepsilon : \{b, c\}, (E - \varepsilon)^c : \{b\})$. Note that by removing ε from E, the relative value of b over c is lower in the event

[3]Note that whether the decision-maker actually chooses subsequent act b or c is not observable until the second stage, but this does not affect our method as it only relies on the value v which is identifiable from first stage preferences.

$E - \varepsilon$ than in E. We then compare the marginal change of v. Suppose first $\varepsilon \subseteq B \cap E \cap I^c$. Since removing ε from E only decreases the value of b in $E \cap I^c$, the decision-maker will still choose c in $E \cap I^c$. Thus,

$$v(\{b,c\}, E \setminus \varepsilon) = v(\{b\}, E \cap I) + v(\{c\}, E \cap I^c \setminus \varepsilon) \tag{5}$$

$$= v(\{b\}, E \cap I) + v(\{c\}, E \cap I^c) - v(\{c\}, \varepsilon) \tag{6}$$

$$= v(\{b,c\}, E) - v(\{c\}, \varepsilon). \tag{7}$$

Suppose instead $\varepsilon \subseteq B \cap E \cap I$. Since the decision-maker prefers b over c in $E \cap I$, removing ε from $E \cap I$ may cause preference reversal. Thus,

$$v(\{b,c\}, E \setminus \varepsilon) = v(\{b,c\}, E \cap I \setminus \varepsilon) + v(\{c\}, E \cap I^c) \tag{8}$$

$$= \begin{cases} v(\{b,c\}, E) - v(\{b\}, \varepsilon), & \text{if } b \text{ is chosen in } E \cap I \setminus \varepsilon, \\ v(\{c\}, E \setminus \varepsilon), & \text{if } c \text{ is chosen in } E \cap I \setminus \varepsilon. \end{cases} \tag{9}$$

If ε is sufficiently large, then $v(\{b,c\}, E \setminus \varepsilon)$ is equal to $v(\{c\}, E \setminus \varepsilon)$, allowing us to distinguish the case $\varepsilon \subseteq E \cap I$ from $\varepsilon \subseteq E \cap I^c$.

The following proposition shows that

Proposition 2. *Assume Axiom 7 and 8 hold. Suppose $\succsim$ can be represented by $\sum_E \max_{b \in m} v_E(\{b\})$. Let b and c be two subsequent acts. Define*

$$B = \{s \in \mathcal{S} : b(s) \succ c(s)\} \text{ and } C = \{s \in \mathcal{S} : c(s) \succ b(s)\}.$$

Suppose

1. $B \cup C = \mathcal{S}$,
2. $\{b,c\}_E\{b\} \succ \{b\}_E\{b\}$ *and* $\{b,c\}_E\{b\} \succ \{c\}_E\{b\}$.

Then there exists $I \in \mathcal{H}$ such that

$$\{b,c\} \sim_{E \cap I} \{b\} \text{ and } \{b,c\} \sim_{E \cap I^c} \{c\}.$$

and for any nonnull $\varepsilon \subseteq B \cap E$, we have that $\varepsilon \cap I$ is null if and only if

$$v(\{b,c\}, E \setminus \varepsilon) = v(\{b,c\}, E) - v(\{c\}, \varepsilon). \tag{10}$$

Proof. We shall prove this propostition in two steps.

Step 1. Show $v(\{b,c\},E) \neq v(\{c\},E)$ and $v(\{b\},\varepsilon) \neq v(\{c\},\varepsilon)$.

Since $\{b,c\}_E\{b\} \succ \{c\}_E\{b\}$, we have $v(\{b,c\},E) > v(\{c\},E)$. The second inequality directly follows from the assumption that ε is a non-null subset of $B \cap E$.

Step 2. Show if $\varepsilon \subsetneq B \cap E \cap I^c$, then the desired equality does not hold.

Let $\varepsilon = \varepsilon' \cup \varepsilon''$, where $\varepsilon' = \varepsilon \cap I$ and $\varepsilon'' = \varepsilon \cap I^c$. Then

$$
\begin{aligned}
&v(\{b,c\},E \setminus \varepsilon) \\
&= v(\{b,c\},E \cap I \setminus \varepsilon') + v(\{b,c\},E \cap I^c \setminus \varepsilon'') \\
&= \left\{ \begin{array}{l} v(\{b,c\},E \cap I) - v(\{b\},\varepsilon') \\ v(\{c\},E \cap I) - v(\{c\},\varepsilon') \end{array} \right\} + v(\{c\},E \cap I^c) - v(\{c\},\varepsilon'') \\
&= \begin{cases} v(\{b,c\},E) - v(\{b\},\varepsilon') - v(\{c\},\varepsilon'') & \text{if } b \text{ is chosen in } E \cap I \setminus \varepsilon'; \\ v(\{c\},E) - v(\{c\},\varepsilon) & \text{if } c \text{ is chosen in } E \cap I \setminus \varepsilon'. \end{cases}
\end{aligned}
$$

Neither of the expressions is equal to $v(\{b,c\},E) - v(\{c\},\varepsilon)$. □

Similarly, we can identify the region C in the upper left block, i.e., $C \cap (E \cap I)$. Specifically,

$$C \cap (E \cap I) = \bigcup \{\varepsilon \subseteq C \cap E \mid v(\{b,c\},E - \varepsilon) = v(\{b,c\},E) - v(\{b\},\varepsilon)\}.$$

Finally, we can use the above sets to recover $E \cap I$ and $E \cap I^c$.

It is of interest that whether one can get a different result if he/she considers menus consisting of more than two subsequent acts. However, with the following lemma we argue that it is sufficient to focus on all binary menus to identify the hidden information partition. To state the result, we generalize the idea of $\mathcal{P}_2$: Given a set of subsequent acts $\{b^1, b^2, \ldots, b^n\}$, let $\mathcal{P}_i$ denote the coarsest partition that is finer than all hidden information partitions generated by a menu consisting of any combination of i subsequent acts.

Lemma 1. *$\mathcal{P}_2$ is finer than $\mathcal{P}_i$ for $i = 3, 4, \ldots, n$.*

Proof. This follows straightforward from the fact that if a hidden information set is not identifiable from the binary choice between b and c, the binary choice between c and d, and the binary choice between b and d, then it cannot be identified from the choice between $\{b, c, d\}$ as the partition of the state space on which b, c, d are optimal from $\{b, c, d\}$ is a coarsening of the common refinement of the partitions on which b, c are optimal from $\{b, c\}$, c, d are optimal from $\{c, d\}$, and b, d are optimal from $\{b, d\}$. Since only border-crossing information sets are identifiable, it follows that $\mathcal{P}_2$ refines $\mathcal{P}_3$. A similar argument holds for choices from larger menus. □

From the above lemma follows that $\mathcal{P}_2$ contains all decision-relevant hidden information of the decision maker. It follows that Proposition 2 uniquely characterizes all decision-relevant hidden information.

8 Conclusion

In this paper we have analyzed preference for flexibility and knowledge with an objective state space. This is in contrast to previous analyses in which the uncertainty about the state space inducing the preference for flexibility is allowed to be completely unrelated to the objective state space. In our model, from information preferences we can identify which information the decision maker expects to gain that induces a preference for flexibility. In a state independent model this allows for separation of utilities, beliefs, and information sigma algebras. Perhaps more surprisingly even in a state-dependent model the hidden information that the decision maker expects to gain is identifiable and only requires the use of binary menus.

Acknowledgements

This work was financially supported by the Center for Research in Econometric Theory and Applications (Grant No. 109-L900-203) from the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan, and by the Ministry of Science and Technology (MOST), Taiwan, under Grant No. 109-2410-H-002-158.

A Further Decision Models

We introduce several decision models capturing various aspects of behavior when agents can choose between acts that may offer both flexibility and knowledge in addition to uncertain outcomes.

In the previous model, we assume that the decision-maker can only receive information through information acts with an information partition known to the analyst. Here, we instead assume that the decision-maker gathers her own information prior to the second stage. This hidden information is unknown to the analyst and hence the decision-maker's preferences violate the indirect utility property from the analyst's perspective. For example, suppose that the decision-maker has the hidden information (partition) $\{E^1, E^2\}$ and assigns equal probability to these events. Assume her utility of b^1 and b^2 are 20 and 0 respectively if E^1 occurs while 0 and 10 if E^2 occurs. Then the decision-maker will report

$$\{b^1\}_{\mathcal{S}} \succ \{b^2\}_{\mathcal{S}} \quad \text{and} \quad \{b^1, b^2\}_{\mathcal{S}} \succ \{b^1\}_{\mathcal{S}},$$

which violates the indirect utility property from the analyst's perspective. The following representation allows for almost arbitrary knowledge and flexibility preferences. In particular, the value of knowing an event may depend on subsequent acts available and the value of the final outcomes may be state-dependent.

Definition 10 (Event Dependent Indirect Utility)**.** $\succsim$ on the set $\mathcal{A}$ defined on events $\mathcal{E}$ has an event dependent indirect utility representation if there exist functions $U : \mathcal{A} \to \mathbb{R}$ and $v : \mathcal{B} \times \mathcal{E} \to \mathbb{R}$ such that $U(a) \geq U(a')$ if and only if $a \succsim a'$ and

$$U(a) = \sum_{E \in \iota(a)} \max_{b \in a(E)} v(b, E). \tag{11}$$

Second, we characterize a representation in which the value of outcomes are state-independent but the value of subsequent acts may depend on the event that obtains. We also exclude any intrinsic preference for knowing whether an event obtains. In this case, beliefs are identifiable and can be represented by a quantitative probability.

If the decision maker has no intrinsic value of knowledge, the decision maker may still prefer to have more knowledge over which states obtains in order to make a better choice from the available menu. In this case, the decision maker has an expecte indirect utility.

Definition 11 (Expected Indirect Utility)**.** $\succsim$ on the set $\mathcal{A}$ defined on events $\mathcal{E}$ has an expected indirect utility representation if there exist functions $U : \mathcal{A} \to \mathbb{R}$ and $u : \mathcal{X} \to \mathbb{R}$ unique up to affine transformations, and a unique probability measure $\mu : \mathcal{E} \to \mathbb{R}$ such that $U(a) \geq U(a')$ if and only if $a \succsim a'$ and

$$U(a) = \sum_{E \in \iota(a)} \mu(E) \max_{b \in a(E)} \int_{s \in E} u\big(b(s)\big) \, \mathrm{d}\mu | E. \tag{12}$$

The crucial difference to the expected indirect utility of Nehring (1999) is that in our model the information that the decision maker expects to receive is objectively given by $\iota(a)$. In our last model, we allow for both subjective information and objective information to play a role. However, the subjective information remains embedded in our objective state space.

We define the preferences of a decision maker who is consequentialist and expects to receive information before choosing the subsequent acts:

Definition 12 (Hidden Information Utility)**.** $\succsim$ on the set $\mathcal{A}$ defined on events $\mathcal{E}$ has a hidden information utility representation if there exist functions $U : \mathcal{A} \to \mathbb{R}$, $u : \mathcal{X} \to \mathbb{R}$ unique up to affine transformations, a unique probability measure $\mu : \mathcal{E} \to \mathbb{R}$, and a hidden information sigma algebra $\mathcal{H}$ such that $U(a) \geq U(a')$ if and only if $a \succsim a'$ and

$$U(a) = \sum_{E \in \iota(a)} \mu(E) \int_{E} \max_{b \in a(E)} w_{b_E} \, \mathrm{d}\mu^{\mathcal{H}|E}. \tag{13}$$

where $w_{b_E}(H) = \int_H u \circ b_E \, \mathrm{d}\mu^{\mathcal{E}|H}$.

We therefore have that in this representation the decision maker first forms an expectation across information sets E provided by an information act. Then, on every information set E the decision maker maximizes across the menu $a(E)$ using the full information from $\mathcal{H}$ to condition the chosen subsequent acts. The maximization objective in this process is given by an expected utility v_b. It is noteworthy that $\max_{b \in a(E)} v_b$ is always measurable since subsequent acts are simple.

B State-Independent Case

An event E is *nonnull* if for some outcomes α and β, we have $\{\alpha\}_E\{\beta\} \succ \{\beta\}_E\{\beta\}$. An event is *null* if it is not nonnull.

Let $\mathcal{X}$ be a set of final *outcomes*. We shall use greek letters to denote generic elements in $\mathcal{X}$. A *subsequent act* b is an $\mathcal{E}$-measurable function

$b : \mathcal{S} \to \mathcal{X}$. If b is a constant function, then we say b is a *constant subsequent act*. The set of all subsequent acts with finite support is denoted $\mathcal{F}_{\mathcal{S}}$.

The next axiom is again a standard axiom in the Savage framework.

Axiom 9 (Nontriviality)**.** There are $\alpha, \beta \in \mathcal{X}$ such that $\alpha \succ \beta$.

Some additional notation will be convenient when menus have this additional structure. If m only consists of constant subsequent acts, then m is called a *constant menu*. Given a subsequent act b, it is useful to consider b restricted on an event E. A *conditional subsequent act* b_E is the function $b_E : E \to \mathcal{X}$ such that $b_E(s) = b(s)$ for all $s \in E$. Similarly, $m_E := \{b_E : b \in m\}$ is called a *conditional menu*. The set of all menus restriced on E is denoted $\mathcal{M}_E$ and the collection of all conditional menus is denoted $\mathcal{M}$, i.e., $\mathcal{M} = \bigcup_{E \in \mathcal{E}} \mathcal{M}_E$.

Definition 13 (Convergence of Subsequent Acts)**.** A sequence (b^k) of subsequent acts converges in subjective belief to b, denoted $b^k \rightsquigarrow b$, if for all $\alpha \in supp(b)$, $(b^k)^{-1}(\alpha) \rightsquigarrow b^{-1}(\alpha)$.

Thus, a sequence of subsequent acts converges to an act if the sequence of events on which an outcome is acquired converges to the event in which the outcome arises on the act the sequence converges to.

Definition 14 (Convergence of Menus)**.** A sequence (m^k) of menus converges in subjective belief to m, denoted $m^k \rightsquigarrow m$, if the following two conditions hold:

1. For all subsequent acts $b \in m$, there exists a sequence (b^k) such that $b^k \in m^k$ for all k and $b^k \rightsquigarrow b$.
2. If there exists a sequence (b^i) such that $b^i \in m^{k_i}$ for all i, where m^{k_i} is a subsequence of m^k and $b^i \rightsquigarrow b$, then $b \in m$.

A menu of subsequent acts converges if the set of acts in the menu is not too large and not too small. If we can pick for every menu in the sequence a subsequent act such that this sequence of subsequent acts converges, then the subsequent act that the sequence converges to must be in the menu. On the converse, for every act in the menu, we need to be able to find a convergent subsequence of acts by picking from every menu in the sequence of menus.

Definition 15 (Convergence of Information Acts)**.** A sequence (a^k) of information acts converges in subjective belief to a, denoted $a^k \rightsquigarrow a$, if the following two conditions hold:

1. $\iota(a^k) \rightsquigarrow \iota(a)$.
2. $a^k(E_i^k) \rightsquigarrow a(E_i)$ for each $E^i \in \iota(a)$, where (E_i^k) is the sequence that converges to E_i.

For the state-independent model, we impose the following monotonicity and likelihood outcome independence axioms.

Axiom 10 (Monotonicity)**.** For all nonnull $E \in \mathcal{E}$, all constant menus m, n and all $f \in \mathcal{A}$,

$$m \succsim n \Longleftrightarrow m_E f \succsim n_E f. \tag{14}$$

Monotonicity captures that if a menu of constant acts is preferred to another menu of constant acts, then it is also preferred conditionally on any event E. This means that both the final outcomes and the flexibility of having the choice between different outcomes are valued state-independently. By comparing singleton constant menus, we can obtain a preference relation over outcomes. For notational convenience, we simply denote $\alpha \succsim \beta$ to express that outcome α is at least as good as β.

Axiom 11 (Likelihood Outcome Independence)**.** For all events $E, F \in \mathcal{E}$ and all constant menus m, n, m', n' with $m \succ n$ and $m' \succ n'$,

$$m_E n \succsim m_F n \Longleftrightarrow m'_E n' \succsim m'_F n'. \tag{15}$$

Likelihood outcome independence guarantees the existence of a *likelihood relation* $\succsim^*$ on the set of events $\mathcal{E}$. Event E is defined to be at least as likely as F, denoted by $E \succsim^* F$, if $\{\alpha\}_E\{\beta\} \succsim \{\alpha\}_F\{\beta\}$ for some outcomes $\alpha, \beta \in \mathcal{X}$ such that $\alpha \succ \beta$.

Likelihood outcome independence only applies to singleton menus. It is thus formulated here in the same form as in Savage (1954) and does not make use of the additional menu structure.[4]

Axioms 1-6 together with the instrumental knowledge property guarantee that for singleton menus, the axioms of Savage (1954) apply and an expected utility representation exists. This is because combining the instrumental knowledge property with the monotonicity axiom, we obtain

$$\begin{aligned} & \{\alpha_E b\} \succsim \{\beta_E b\} \\ \Leftrightarrow\ & \{\alpha\}_E\{b\} \succsim \{\beta\}_E\{b\} \\ \Leftrightarrow\ & \qquad \{\alpha\} \succsim \{\beta\}. \end{aligned}$$

[4]However, the indirect utility property over menus together with likelihood outcome independence guarantee that likelihood judgments are also consistent when elicited using menus of constant subsequent acts.

Axiom 12 (Consistency of Hidden Knowledge)**.** Let $\alpha, \beta \in \mathcal{X}$ with $\{\alpha\} \succ \{\beta\}$. If an event I and the menu $n = \{\alpha_I\beta, \beta_I\alpha\}$ satisfy $n \sim n_I n$, then for all $E \in \mathcal{E}$ and all menus m, $m \sim_E m_I m$.

Consistency of hidden knowledge ensures that if a decision maker does not value information whether I or I^c occurs for a menu in which this information should be useful, then the decision maker never attaches value to this information.

Definition 16 (Hidden Identified Set)**.** An event I is called a hidden identified set if I satisfies for some $\{\alpha\} \succ \{\beta\}$ and $m = \{\alpha_I\beta, \beta_I\alpha\}$, $m \sim m_I m$.

Axiom 13 (Hidden Indirect Utility Property)**.** Let m and n be two menus. If $m \cup n \succ_E m$ and $m \cup n \succ_E n$, then there exist a hidden identified set I such that $m \cup n \sim_{E \cap I} m$ and $m \cup n \sim_{E \cap I^c} n$.

According to the hidden indirect utility property, whenever a decision maker has a strict preference for flexibility on event E for choosing from menu $m \cup n$ rather than either of the two menus, then there exists some information that the decision maker expects to receive and this information will make a choice from m optimal in some states of the world and a choice from n optimal in other states of the world. The hidden indirect utility property prevents preference for flexibility to be due to a subjective state space. Thus, the objective state space captures all uncertainty relevant for making a choice from any menu.

Next, we provide a representation in which the outcomes of acts are evaluated state-independently and in which probabilistic beliefs are identified. In this theorem we assume both the indirect utility property for subsequent acts and the instrumental knowledge property to obtain that the decision maker maximizes a utility over subsequent conditional acts at every information set. The utility over subsequent acts may or may not be consistent with expected utility maximization and may capture a variety of behavioral biases. While we focus on expected utility maximization in this paper, this representation may be a useful starting point for decision models that capture a greater variety of behavior.

Theorem 4. *The following statements are equivalent:*

1. $\succsim$ *satisfy Axioms 1-6, Axiom* **??***, and Axiom 7.*

2. $\succsim$ *has an indirect expected utility representation.*

The proof in appendix F starts by showing that $\succsim^*$ is a qualitative probability and by the richness of states we obtain a representation of this relation by a unique quantitative probability. Factoring out this quantitative probability from the state-dependent representation obtained in Theorem **??**, we can then use the indirect utility property to obtain utility maximization across subsequent acts. This in turn allows us to express the utility of every information act as an information act that yields a singleton (the optimal subsequent act) on every information set. Using the instrumental knowledge property we can additively decompose the utility of a subsequent act given an event E into the outcomes. This is straightforward to see by observing that on the set of information acts only yielding singleton menus our axioms imply the axioms of Savage (1954).

C Proof of Theorem 1

We proceed in three major steps. First we prove that the order topology on conditional acts is connected. This allows us to use standard additive representation theorems to obtain for every fixed information partition a utility that is additively separable across elements of the information partition. Finally, we show that for every event the additive component is independent of the remainder of the partition.

Definition 17.

1. Let $E \triangle F$ denote the symmetric difference of two sets E and F, i.e.,

$$E \triangle F := (E \setminus F) \cup (F \setminus E).$$

2. Let $\hat{\mu}$ be an atomless probability measure on $\mathcal{E}$ satisfying $\hat{\mu}(E) = 0$ if and only if E is null.

3. Let Σ denote a subset of $\mathcal{E}$ satisfying the following three conditions: An issue here. Given an ideal, is there always exist a consistent probability measure? How to say $\mathcal{E}$ is atomless without specifying the measure or qualitative probability?

   (a) $\emptyset, \mathcal{S} \in \Sigma$;

   (b) for each $p \in [0, 1]$, exactly one $F \in \Sigma$ satisfies $\hat{\mu}(F) = p$;

(c) for any two events F^i, $F^j \in \Sigma$, either $F^i \subseteq F^j$ or $F^j \subseteq F^i$.

4. Given two information acts f, g and two events F, $G \in \Sigma$ with $F \subseteq G$, define
$$\mathcal{I}(f,g,F,G) := \{f_E g : E \in \Sigma \text{ and } F \subseteq E \subseteq G\}.$$
If $F = \emptyset$ and $G = \mathcal{S}$, then F and G will be dropped and we simply write $\mathcal{I}(f,g)$.

Lemma 2. *Assume Axiom 1, 2, and 4 hold. Let $E \in \mathcal{E}$ be an event. Let f, g be two information acts and F, $G \in \Sigma$ be two events. Then $\mathcal{I}(f,g,F,G)$ is connected if endowed with the order topology of $\succsim_E$.*

Proof. For notational convenience, denote $\mathcal{I}(f,g,F,G)$ by $\mathcal{I}$. Consider the set
$$I := \left\{\hat{\mu}(H) : H \in \Sigma,\ F \subseteq H \subseteq G, \text{ and } a \succsim_E f_H g\right\},$$
where $a \in \mathcal{I}$. Pick a convergent sequence (z^k) in this set with $\lim z^k = z^*$. Let H^k and H^* denote the events in Σ satisfying $\hat{\mu}(H^k) = z^k$ and $\hat{\mu}(H^*) = z^*$. We claim that $H^k \rightsquigarrow H^*$. Pick a nonnull event $J \subseteq H^*$ and assume by contradiction that (E1) does not hold. Then there exists a subsequence (H^p) such that $J \cap H^p = \emptyset$ for all p, which also implies $H^p \subseteq H^*$. Hence, $\hat{\mu}(J) + \hat{\mu}(H^p) \leq \hat{\mu}(H^*)$, a contradiction. Assume by contradiction that (E2) does not hold. Construct the subsequence (H^q) accordingly, i.e., there exists nonnull event K such that $K \subsetneq H^q$ for all q and $K \cap H^* = \emptyset$. This implies $H \subseteq H^q$ and $\hat{\mu}(K) + \hat{\mu}(H^*) \leq \hat{\mu}(H^q)$ for all q, a contradiction. Thus, $H^k \rightsquigarrow H^*$.

Pick $h \in \mathcal{A}$. Observe that $(f_{H^k} g)_E h \rightsquigarrow (f_{H^*} g)_E h$. Since $a \succsim_E f_{H^k} g$ for all k, we have $a_E h \succsim (f_{H^k} g)_E h$ for all k. By Axioms 4, $a_E h \succsim (f_{H^*} g)_E h$. That is, $z^* = \hat{\mu}(H^*) \in I$. Hence, I is closed. Similarly,
$$\left\{\hat{\mu}(H) : H \in \Sigma,\ F \subseteq H \subseteq G, \text{ and } f_H g \succsim_E a\right\}$$
is closed for all $a \in \mathcal{I}$. By taking complement and intersection, the following three types of sets are open for all $\hat{a}$, $\check{a} \in \mathcal{I}$:

1. $\left\{\hat{\mu}(H) : H \in \Sigma,\ F \subseteq H \subseteq G, \text{ and } \hat{a} \succ_E f_H g\right\}$;
2. $\left\{\hat{\mu}(H) : H \in \Sigma,\ F \subseteq H \subseteq G, \text{ and } f_H g \succ_E \check{a}\right\}$; and
3. $\left\{\hat{\mu}(H) : H \in \Sigma,\ F \subseteq H \subseteq G, \text{ and } \hat{a} \succ_E f_H g \succ_E \check{a}\right\}$.

Assume by contradiction that $\mathcal{I}$ is not connected. Then there exist two nonempty open sets U, V such that $U \cap V = \emptyset$ and $U \cup V = \mathcal{I}$. Note that

$$\{h \in \mathcal{I} : \hat{a} \succ_E h\}, \ \{h \in \mathcal{I} : h \succ_E \check{a}\}, \text{ and } \{h \in \mathcal{I} : \hat{a} \succ_E h \succ_E \check{a}\}$$

for all $\hat{a}$, $\check{a} \in \mathcal{I}$ form a base for the order topology of $\succsim_E$. Hence, both U and V are unions of such open intervals ard rays. As have already shown, openness is preserved for those three types of sets if representing in the form of $\hat{\mu}(H)$. Moreover, any union of open sets is open. Thus,

$$\{\hat{\mu}(H) : f_H g \in U\} \text{ and } \{\hat{\mu}(H) : f_H g \in V\}$$

are open and nonempty. These two sets form a partiton of $[\hat{\mu}(F), \hat{\mu}(G)]$, contradicting to the fact that $[\hat{\mu}(F), \hat{\mu}(G)]$ is connected. □

Lemma 3. *Assume Axiom 1, 2, and 4 hold. Let $E \in \mathcal{E}$ be an event. The set $\mathcal{A}$ is connected if endowed with the order topology of $\succsim_E$.*

Proof. If E is null, then $\mathcal{A}$ is topologically a point and hence is connected. From now on, assume E is nonnull.

To show $\mathcal{A}$ is connected, it is equivalent to show $\mathcal{A}$, endowed with the order topology of $\succsim_E$, is a linear continuum. That is, we shall prove:

1. For any two conditional acts $f, g \in \mathcal{A}$ with $f \succ_E g$, there exists $h \in \mathcal{A}$ such that $f \succ_E h \succ_E g$.

2. $\mathcal{A}$ has the least upper bound property with respect to $\succsim_E$.

For the first statement, note that $\mathcal{I}(f, g)$ is connected and hence there must exist $h \in \mathcal{I}(f, g) \subseteq \mathcal{A}$ such that $f \succ_E h \succ_E g$.

For the second statement, take a bounded subset A of $\mathcal{A}$. Pick $f \in \mathcal{A}$ and let $\hat{h}$ be a bound. Denote $\mathcal{I}(f, \hat{h})$ by $\mathcal{I}$. Since $\mathcal{I}$ is connected, for each $a \in A$ satisfying $a \succsim_E f$, there exists $g^a \in \mathcal{I}$ such that $g^a \sim_E a$. Otherwise, $\{h \in \mathcal{I} : h \succ_E a\}$ and $\{h \in \mathcal{I} : a \succ_E h\}$ are two nonempty open sets which together form a partition of $\mathcal{I}$, a contradiction. Let B be the collection of all corresponding g^a. Then B is a nonempty subset of $\mathcal{I}$ and is bounded by $\hat{h} \in \mathcal{I}$. Since $\mathcal{I}$ is a linear continuum, B has a least upper bound g^*. We claim that g^* is also a least upper bound of A. First, for each $a \in A$, if $f \succsim_E a$, then $g^* \succsim_E f \succsim_E a$. If $a \succsim_E f$, then $g^* \succsim_E g^a \sim_E a$. Hence, g^* is an upper bound of A. Next, suppose there exists another upper bound h of A but $g^* \succ_E h$. Clearly, $\hat{h} \succsim_E g^* \succ_E h \succsim_E f$. Hence, there exists $g^h \in \mathcal{I}$ such that $g^h \sim_E h$. Since g^* is a least upper bound of B and $g^* \succ_E g^h$, there must exist $g^a \in B$ such that $g^a \succ_E g^h$. By construction, there exists $a \in A$ such that $a \sim_E g^a \succ_E g^h \sim_E h$, a contradictin.

$\square$

Lemma 4. *Assume Axiom 1, 2, and 4 hold. $\succsim$ is continuous with respect to the product topology.*

Proof. (sketch) Show convergence in the product topology (of $\succsim_E$ and $\succsim_F$) implies convergence in the order topology of $\succsim_G$, where $G = E \cup F$. Why does this work? Continuity means that the weak upper contour set is closed with respect to the product topology. If $f^n \succsim g$ and $f^n \xrightarrow{PT} f$, then $f \succsim g$, which is consistent with convergence in the order topology of $\succsim_G$. Alternatively, we need to show the product topology is finer than the order topology of $\succsim_G$ since every strict upper contour set (open with respect to the order topology) should be open with respect to the product topology.

Assume by contradiction that $(f_E^k f_F^k f^k)$ does not converge in the order topology of $\succsim_G$ to $f_E f_F f$. Then there exists a subsequence $(f_E^p f_F^p f^p)$ and two information acts $\hat{a}$, $\check{a}$ such that either $f_E^p f_F^p f^p \succsim_G \hat{a} \succ_G f_E f_F f$ for all p, or $f_E f_F f \succ_G \check{a} \succsim_G f_E^p f_F^p f^p$ for all p. Without loss of generality, assume the former is the case.

For notational convenience, denote $f_E f_F f$ by f_{EF} and treat other indexed information acts similarly. Since $f_{EF}^p \xrightarrow[E\times F]{PT} f_{EF}$, we have $f_{EF}^p \xrightarrow[E]{OT} f_{EF}$ and $f_{EF}^p \xrightarrow[F]{OT} f_{EF}$. Consider $f_{EF}^p \xrightarrow[E]{OT} f_{EF}$ first. Pick a monotone subsequence (f_{EF}^q) of (f_{EF}^p) with respect to $\succsim_E$. Let

$$\mu(H) = \inf\{\mu(H') : H' \in \Sigma \text{ and } (f_{EF}^1)_{H'}(f_{EF}) \sim_E f_{EF}\}.$$

Hence, there exists a decreasing sequence $(H^k) \subseteq \Sigma$ such that $(f_{EF}^1)_{H^k}(f_{EF}) \sim_E f_{EF}$ for all k and $\mu(H^k) \to \mu(H)$. By Axiom 4, $(f_{EF}^1)_H(f_{EF}) \sim_E f_{EF}$. For each $q > 1$, find $g^q \in \mathcal{I}(f_{EF}^1, f_{EF}, H^{q-1}, H)$ such that $g^q \sim_E f_{EF}^q$, where H^{q-1} satisfies $g^{q-1} = (f_{EF}^1)_{H^{q-1}} H$. Pick a monotone subsequence (f_{EF}^r) of (f_{EF}^q) with respect to $\succsim_F$. Let

$$\mu(K) = \inf\{\mu(K') : K' \in \Sigma \text{ and } (f_{EF}^1)_{K'}(f_{EF}) \sim_F f_{EF}\}$$

and find $h^r \in \mathcal{I}(f_{EF}^{r-1}, f_{EF}, K)$ such that $h^r \sim_F f_{EF}^r$.

By Axiom 2, for all r

$$g_E^r h_F^r f^r \sim g_E^r f_F^r f^r \sim f_E^r f_F^r f^r.$$

Hence, $g^r_E h^r_F \sim_G f^r_E f^r_F \succsim_G \hat{a}$.

Since $g^r_E h^r_F \rightsquigarrow g_E h_F \succsim_G f_E f_F$, by Axiom 4,

$$f_E f_F \sim_G g_E h_F \succsim_G \hat{a} \succ_G f_E f_F,$$

a contradiction. □

Definition 18. Given an information partition $\mathcal{P}$ of $\mathcal{S}$, the collection of all information acts that have the same information partition as $\mathcal{P}$ is denoted by $\mathcal{A}(\mathcal{P})$. Namely,

$$\mathcal{A}(\mathcal{P}) = \{a \in \mathcal{A} : \iota(a) = \mathcal{P}\}.$$

Lemma 5. *Fix a partition $\mathcal{P}$ such that the number of nonnull events is at least three. Let $\succsim$ be restricted to the set of acts a with an information partition $\iota(a)$ at least as fine as $\mathcal{P}$. Assume Axiom 1, 2, and 4 hold. Then $\succsim$ has a representation*

$$U(a) = \sum_{E \in \mathcal{P}} v_E(a_E; \mathcal{P}).$$

Moreover, v_E is continuous and is unique up to eventwise jointly positive linear transformations and eventwise separate additive transformations.

Proof. We have already shown that the set of conditional acts given events is connected. By continuity, the sure-thing principle and theorem III.4.1 in Wakker (1989) the result follows. □

The last step is to show that the representation is partition independent. Consider any partition $\mathcal{P}$ and a refinement $\mathcal{P}'$. Notice that the additive representation across $\mathcal{P}'$ holds on a subset of the acts for which we have an additive representation across $\mathcal{P}$. On this subset, it follows from the uniqueness properties of both representations, that these are affine transformations of another. Without loss of generality, assume that this is the identity transformation. Thus, for any $E^1, \ldots, E^K \in \mathcal{P}'$ such that their union is $E \in \mathcal{P}$, we have that $\sum_k v_{E^k}(a_{E^k}, \mathcal{P}') = v_E(a_E, \mathcal{P})$. In a similar manner, it can be shown that for any two partitions $\mathcal{P}$ and $\mathcal{P}'$ containing event E, we can ensure $v_E(\cdot, \mathcal{P}) = v_E(\cdot, \mathcal{P}')$. We therefore obtain a representation $\sum_{E \in \iota(a)} v_E(a(E))$.

The uniqueness properties follow from the uniqueness properties of the additive representations which are unique up to joint linear and separate additive representations. However, the additive transformations $\eta : \mathcal{E} \to \mathbb{R}$ are restricted to fulfill $v_E(a_E) + \eta(E) + v_F(a_F) + \eta(F) = v_{E \cup F}(a_E) + \eta(E \cup F)$ and thus η must be an additive set function.

D Proof of Proposition ??

Proof. Necessity of the axioms is straightforward. To prove sufficiency, we first define the optimal choice from m.

Definition 19. Given an arbitrarily enumerated menu $m = \{f^1, \dots, f^n\}$ and an event E $f^i = \max_{\succsim_E} m$ if $f^i \in m$ and $\{f^i\} \succsim_E \{f^j\}$ for all $f^j \in m$. If there are more than one element satisfying the condition, then we pick the one with the greatest index number.

Next, we provide a lemma that shows that every menu is equally preferred as its best element.

Lemma 6. *If Axiom **??** holds, then for any menu m, $m \sim_E \max_{\succsim} m$.*

Proof. Let $m = \{f_1, \dots, f_k\}$ be enumerated such that $f_i \succsim_E f_j$ if $i \geq j$. Then by the indirect utility property, $f_2 \sim_E \{f_1, f_2\}$. Moreover, if $f_i \sim \{f_1, \dots, f_i\}$, then since $f_{i+1} \succsim_E f_i$, by transitivity and the indirect utility property $f_{i+1} \sim_E \{f_1, \dots, f_{i+1}\}$. By induction follows that $\max_{\succsim_E} m = f_k \sim_E m$. □

Finally, we prove sufficiency. Starting from an additive representation across information sets, for any event $E \in \iota(a)$, by Lemma 6, $m \sim_E \max_{\succsim_E} m$ and by Axiom 2 $v(E, m) = v(E, \{\max_{\succsim_E} m\}) = \max f \in mv(E, \{f\})$. □

E Proof of Theorem ??

Before we state the main characterization result, we make two key observations that are crucial for recovering an information sigma algebra for the hidden information.

Lemma 7. *Let $\succsim$ be represented by $U(a) = \sum_{E \in \iota(a)} v(a(E), E)$ with $v(m, E) = v(m', E)$ whenever m and m' agree on E and fulfill Axiom 12. If I and J are hidden identified sets, then $I \cap J$ is also a hidden identified set.*

Proof. Since I and J are hidden identified sets, we have

$$v(m, I) + v(m, I^c) = v(m, \mathcal{S}) = v(m, J) + v(m, J^c).$$

Let $m = \{\alpha_{I\cap J}\beta_{I\cap J^c}\gamma, \beta_{I\cap J}\alpha_{I\cap J^c}\gamma\}$. Then

$$\begin{aligned}
v(m, I) &= v(\{\alpha_{I\cap J}\beta, \beta_{I\cap J}\alpha\}, I) \\
v(m, I^c) &= v(\{\gamma\}, I^c) = v(\{\gamma\}, I^c \cap J) + v(\{\gamma\}, I^c \cap J^c) \\
v(m, J) &= v(\{\alpha, \beta\}, I \cap J) + v(\{\gamma\}, I^c \cap J) \\
v(m, J^c) &= v(\{\beta, \alpha\}, I \cap J^c) + v(\{\gamma\}, I^c \cap J^c).
\end{aligned}$$

Hence,

$$\begin{aligned}
v(\{\alpha_{I\cap J}\beta, \beta_{I\cap J}\alpha\}, I) &= v(m, I) = v(m, J) + v(m, J^c) - v(m, I^c) \\
&= v(\{\alpha, \beta\}, I \cap J) + v(\{\beta, \alpha\}, I \cap J^c).
\end{aligned}$$

Finally,

$$\begin{aligned}
v(\{\alpha_{I\cap J}\beta, \beta_{I\cap J}\alpha\}, \mathcal{S}) &= v(\{\alpha_{I\cap J}\beta, \beta_{I\cap J}\alpha\}, I) + v(\{\alpha_{I\cap J}\beta, \beta_{I\cap J}\alpha\}, I^c) \\
&= v(\{\alpha, \beta\}, I \cap J) + v(\{\beta, \alpha\}, I \cap J^c) + v(\{\beta, \alpha\}, I^c) \\
&= v(\{\alpha, \beta\}, I \cap J) + v(\{\beta, \alpha\}, I \cap J^c) \\
&= v(\{\alpha_{I\cap J}\beta, \beta_{I\cap J}\alpha\}, I \cap J) + v(\{\alpha_{I\cap J}\beta, \beta_{I\cap J}\alpha\}, I \cap J^c).
\end{aligned}$$

□

Lemma 8. *Let* $\succsim$ *be represented by* $\sum_E v(a(E), E)$ *with* $v(m, E) = v(m', E)$ *whenever* m *and* m' *agree on* E *and fulfill Axiom 4 and 12. If* I^1, I^2, … *are hidden identified sets, then* $I^* = \bigcap_{i=1}^{\infty} I^i$ *is also a hidden identified set.*

Proof. Pick $\alpha, \beta \in \mathcal{X}$ such that $\alpha \succ \beta$. Let $J^p = \bigcap_{i=1}^{p} I^i$. Let $m^p = \{\alpha_{J^p}\beta, \beta_{J^p}\alpha\}$. Since each J^p is a hidden identified set, we have

$$m^p \sim m^p_{J^p} m^p \text{ for all } p.$$

Let $m^* = \{\alpha_{I^*}\beta, \beta_{I^*}\alpha\}$. Since $J^p \searrow I^*$, we have $J^p \rightsquigarrow I^*$ and $m^p \rightsquigarrow m^*$, which implies

$$\begin{gathered}
m^p_{J^p} m^p \rightsquigarrow m^*_{I^*} m^* \text{ and} \\
m^p_{\mathcal{S}} \rightsquigarrow m^*_{\mathcal{S}}.
\end{gathered}$$

By Continuity,

$$m^* \sim m^*_{I^*} m^*.$$

□

Proof. We only prove sufficiency of the axioms. From theorem 1, we obtain a representation $U(f) = \sum_{E \in \iota(f)} v(f(E), E)$. From Lemma 1 and Lemma 2 we know that a (potentially trivial) hidden information sigma algebra exists and that $v(m, \cdot)$ is a measure when restricted to the sub sigma algebra $\mathcal{H}|E$. By the Radon Nikodym Theorem, since v is finite and additive with respect to $\mathcal{H}|E$, and $v(f(E), E) = 0$ if $\mu(E) = 0$, we have that for all $F \in \mathcal{H}|E$, $v(m, F) = \int_F w_F(m, s) d\mu^{\mathcal{H}|F}(s)$ where w is the Radon Nikodym derivative of v with respect to $\mu^{\mathcal{H}|F}$.

We claim that for each E, and for each $I \in \mathcal{H}|E$,

$$v(m, I) = \int_I \max_{b \in m} w_E(\{b\}, s) d\mu^{\mathcal{H}|E}.$$

Let $m = \{b^1, \ldots, b^p\}$, $m^1 = \{b^1\}$, $m^2 = \{b^1, b^2\}$, $\ldots$, and $m^p = m$. The claim is valid when $p = 1$, i.e., $|m| = 1$. Assume by induction that the claim is correct when $p = k$. By the uniqueness of Randon-Nikodym derivative, $\max_{b \in m^k} w_E(\{b\}, s)$ is equal to $w_E(m^k, s)$ up to a $\mu^{\mathcal{H}|E}$-null set.

We distinguish between three possible cases. Suppose first $v(m^{k+1}, E) = v(m^k, E)$. By Axiom 6, for all $I \in \mathcal{H}|E$,

$$v(m^{k+1}, I) \geq \max\left\{v(m^k, I), v(\{b^{k+1}\}, I)\right\}.$$

Assume by contradiction that $v(m^{k+1}, I) > v(m^k, I)$ for some $I \in \mathcal{H}|E$ such that $\mu^{\mathcal{H}|E}(I) > 0$. Then by Axiom 8,

$$\begin{aligned} v(m^{k+1}, E) &= v(m^{k+1}, I) + v(m^{k+1}, E - I) \\ &> v(m^k, I) + v(m^k, E - I) = v(m^k, E), \end{aligned}$$

a contradiction. Hence, for all $I \in \mathcal{H}|E$,

$$v(m^k, I) = v(m^{k+1}, I) \geq v(\{b^{k+1}\}, I).$$

Equivalently, for all $I \in \mathcal{H}|E$,

$$\int_I w_E(m^{k+1}, s) d\mu^{\mathcal{H}|E} = \int_I w_E(m^k, s) d\mu^{\mathcal{H}|E} \geq \int_I w_E(\{b^{k+1}\}, s) d\mu^{\mathcal{H}|E},$$

which implies that

$$w_E(m^{k+1}, s) = w_E(m^k, s) \geq w_E(\{b^{k+1}\}, s) \text{ a.e. with respect to } \mu^{\mathcal{H}|E}.$$

Consequently, the following equalities hold a.e. with respect to $\mu^{\mathcal{H}|E}$:

$$\begin{aligned} w_E(m^{k+1}, s) &= \max\left\{w_E(m^k, s), w_E(\{b^{k+1}\}, s)\right\} \\ &= \max\left\{\max_{b \in m^k} w_E(\{b\}, s), w_E(\{b^{k+1}\}, s)\right\} \\ &= \max_{b \in m^{k+1}} w_E(\{b\}, s), \end{aligned}$$

which implies

$$v(m^{k+1}, E) = \int_E w_E(m^{k+1}, s)\, d\mu^{\mathcal{H}|E} = \int_E \max_{b \in m^{k+1}} w_E\big(\{b\}, s\big)\, d\mu^{\mathcal{H}|E}.$$

The case that $v(m^{k+1}, E) = v(\{b^{k+1}\}, E)$ is analogous. Suppose now $v(m^{k+1}, E) > v(m^k, E)$ and $v(m^{k+1}, E) > v(\{b^{k+1}\}, E)$. By Axiom 8, there exists an hidden identified set H^* such that

$$v(m^{k+1}, E \cap H^*) = v(m^k, E \cap H^*) \qquad \text{and} \tag{16}$$

$$v(m^{k+1}, E \cap \overline{H^*}) = v(\{b^{k+1}\}, E \cap \overline{H^*}). \tag{17}$$

Let $I^* = E \cap H^*$. Note that $I^* \in \mathcal{H}|E$. By Axiom 6, (16), and (17),

$$\begin{aligned} &v\big(m^{k+1}, I\big) = v\big(m^k, I\big) \geq v\big(\{b^{k+1}\}, I\big) \text{ for all } I \in \mathcal{H}|I^* \qquad \text{and} \\ &v\big(m^{k+1}, I\big) = v\big(\{b^{k+1}\}, I\big) \geq v\big(m^k, I\big) \text{ for all } I \in \mathcal{H}|E - I^*. \end{aligned}$$

By the argument similar to the above, we have

$$\begin{aligned} v(m^{k+1}, E) &= v(m^{k+1}, I^*) + v(m^{k+1}, E - I^*) \\ &= \int_{I^*} w_E(m^{k+1}, s)\, d\mu^{\mathcal{H}|E} + \int_{E - I^*} w_E(m^{k+1}, s)\, d\mu^{\mathcal{H}|E} \\ &= \int_{I^*} \max_{b \in m^{k+1}} w_E(\{b\}, s)\, d\mu^{\mathcal{H}|E} + \int_{E - I^*} \max_{b \in m^{k+1}} w_E(\{b\}, s)\, d\mu^{\mathcal{H}|E} \\ &= \int_E \max_{b \in m^{k+1}} w_E(\{b\}, s)\, d\mu^{\mathcal{H}|E}. \end{aligned}$$

Lastly, by the instrumental knowledge property the utility of every subsequent act conditional on an event is additively decomposable into the utility of its outcomes as previously shown for the case without hidden information. □

Lemma 9. *Let $\mathcal{H}_n$ be the closure of $\mathcal{H}_0$ under finite unions and intersections. Then for all disjoint $E, F \in \mathcal{H}_n$, $v(E, m) + v(F, m) = v(E \cup F, m)$.*

Proof. Since E and F are finite unions and intersections of elements of $\mathcal{H}_0$, there exists an enumeration $H_1, \ldots, H_n$ of these elements. For all $k = 1, \ldots, n$, define $\mathcal{P}_k = \{H_1, \overline{H_1}\} \wedge \ldots \wedge \{H_k, \overline{H_k}\}$. $E \cup F$, E, and F can be decomposed into a finite union of elements of the partition $\mathcal{P}_n$.

Lemma 10. *If an event J can be decomposed into elements of $\mathcal{P}_n$, then $v(J, m) = \sum_{P \in \mathcal{P}_n : P \subseteq J} v(P, m)$.*

Proof. Since H_1 is an HIS, $v(J) = v(J \cap H_1, m) + v(J \cap \overline{H_1}, m)$. Since H_k is an HIS, for all elements G of $\mathcal{P}_{k-1}$ and all E, $v(G \cap H_k \cap J, m) + v(G \cap \overline{H_k} \cap J, m) = v(G \cap J, m)$. Note that because $v(\emptyset, m) = 0$, if $G \cap H_k \cap J = \emptyset$ or $G \cap \overline{H_k} \cap J = \emptyset$, the result is trivially true. Otherwise, the result follows from property 4 of an HIS. By induction over k, we then obtain the desired result. □

It then follows from disjointness of E and F that

$$
\begin{aligned}
v(E \cup F, m) &= \sum_{P \in \mathcal{P}_n : P \subseteq E \cup F} v(P, m) \\
&= \sum_{P \in \mathcal{P}_n : P \subseteq E} v(P, m) + \sum_{P \in \mathcal{P}_n : P \subseteq F} v(P, m) \\
&= v(E, m) + v(F, m).
\end{aligned} \tag{18}
$$

□

Let $\mathcal{H}$ be the closure of H_0 under countable unions and intersections. Let $\mathcal{H}|E$ be the conditional sigma algebra of $\mathcal{H}$ given event E.

Lemma 11. *For all m, $v_{E,m}$ defined by $v_{E,m}(F) = v(F \cap E, m)$ is a finite measure on $\mathcal{H}$.*

Proof. First, observe that $v(\mathcal{S}, m)$ is finite. Second, we show that v is sigma-additive.

$$\begin{aligned} v\left(\bigcup_{i=1}^{\infty} F_i, m\right) &= \lim_{n\to\infty} v\left(\bigcup_{i=1}^{\infty} F_i, m\right) \\ &= \lim_{n\to\infty}\left\{ v\left(\bigcup_{i=1}^{n} F_i, m\right) + v\left(\bigcup_{i=n+1}^{\infty} F_i, m\right)\right\} \\ &= \lim_{n\to\infty}\left\{ \sum_{i=1}^{n} v(F_i, m) + v\left(\bigcup_{i=n+1}^{\infty} F_i, m\right)\right\} \\ &= \sum_{i=1}^{\infty} v(F_i, m) + 0. \end{aligned}$$

where the last step follows from monotone continuity of v. □

We now choose an arbitrary probability measure μ on $\mathcal{E}$ such that all null sets have zero probability. Since v is zero on all null sets, it is absolutely continuous with respect to μ. By the Radon-Nikodym theorem, we then have that $v(E, m) = \int_E v_m d\mu^{\mathcal{H}|E}$ for an appropriate density v_m. What is left to do is to determine $v(E, m)$ for all E.

By Axiom 8, for every menu $m \cup n$ and event E there exists an event I such that $m_I n \sim_E m \cup n$. Thus, every $f = m^1_{E_1} \dots m^n \sim \{b^{1,1}_{E_1\cap I_1} b^{1,2}_{E_1\cap I_2} \dots b^{n,k_n}\}$ for arbitrary enumerations $\{b^{i,1}, \dots, b^{i,k_i}\}$ of the elements of m^i. Moreover, $b^{i,k}$ must be optimal on $E_i \cap I_k$ by the Hidden Indirect Utility Property and the Positive Value of Flexibility. Thus, $v(E_i \cap I_k, \{b^{i,k}\}) = v(E_i \cap I_k, m)$ and $v(E_i \cap I_k, \{b^{i,k}\}) \geq v(E_i \cap I_k, \{b'\})$ for all $b' \in m$. It follows that $v(E, m) = \int_E \max_{b\in m} v_b d\mu^{\mathcal{H}|E}$.

F Proof of Theorem 2

We first prove that our likelihood relation is a qualitative probability. Using a result by Villegas (1964), we then show that the likelihood relation can be represented by a quantitative probability. Finally, we factor out the probabilities from the additive representation in Theorem 1 and apply the indirect utility property and the instrumental knowledge property to obtain that the decision maker maximizes across a state independent expected utility over outcomes.

Definition 20 (Qualitative Probability)**.** A binary relation R on $\mathcal{E}$ is a qualitative probability if it fulfills the following conditions:

1. R is a weak order.
2. For all $E \in \mathcal{E}$: $\mathcal{S}RER\emptyset$ and not $\emptyset R\mathcal{S}$.
3. If E, F are disjoint from G, then $ERF \Leftrightarrow E \cup GRF \cup G$.

Lemma 12. *Assume Axiom 1, 2, 10, 9, and 7 hold. Then $\succsim^*$ is a qualitative probability.*

Proof. We check Definition 20 in three steps.

Step 1. $\succsim^*$ is a weak order.

Completeness follows directly from completeness of $\succsim$. Suppose now $E \succsim^* F$ and $F \succsim^* G$. Then for some α, β, γ, $\delta \in \mathcal{X}$ with $\{\alpha\} \succ \{\beta\}$ and $\{\gamma\} \succ \{\delta\}$, we have that $\{\alpha\}_E\{\beta\} \succsim \{\alpha\}_F\{\beta\}$ and $\{\gamma\}_F\{\delta\} \succsim \{\gamma\}_G\{\delta\}$. By Axiom 11, $\{\alpha\}_F\{\beta\} \succsim \{\alpha\}_G\{\beta\}$. By transitivity of $\succsim$, $\{\alpha\}_E\{\beta\} \succsim \{\alpha\}_G\{\beta\}$. Hence, $E \succsim^* G$ and transitivity of $\succsim^*$ holds.

Step 2. For all $E \in \mathcal{E}$, $\mathcal{S} \succsim^* E \succsim^* \emptyset$ and not $\emptyset \succsim^* \mathcal{S}$.

Let $\{\alpha\} \succ \{\beta\}$. If E^c is null, then $\{\alpha\}_E\{\alpha\} \sim \{\alpha\}_E\{\beta\}$. If E^c is nonnull, then by Axiom 10, $\{\alpha\}_E\{\alpha\} \succ \{\alpha\}_E\{\beta\}$. By Axiom 7, $\{\alpha\}_E\{\alpha\} \sim \{\alpha\} = \{\alpha\}_{\mathcal{S}}\{\beta\}$. Thus, in either case, $\{\alpha\}_{\mathcal{S}}\{\beta\} \succsim \{\alpha\}_E\{\beta\}$, which implies $\mathcal{S} \succsim^* E$.

$E \succsim^* \emptyset$ is proven in a similar approach. If E is null, then $\{\alpha\}_E\{\beta\} \sim \{\beta\}_E\{\beta\}$. If E is nonnull, then by Axiom 10, $\{\alpha\}_E\{\beta\} \succ \{\beta\}_E\{\beta\}$. By Axiom 7, $\{\beta\}_E\{\beta\} \sim \{\beta\} = \{\alpha\}_\emptyset\{\beta\}$. Thus, in either case, $\{\alpha\}_E\{\beta\} \succsim \{\alpha\}_\emptyset\{\beta\}$, which implies $E \succsim^* \emptyset$.

Finally, note that $\{\alpha\} = \{\alpha\}_{\mathcal{S}}\{\beta\}$ and $\{\beta\} = \{\alpha\}_\emptyset\{\beta\}$. Hence, $\mathcal{S} \succ^* \emptyset$.

Step 3. If E, F are disjoint from G, then $E \succsim^* F \iff E \cup G \succsim^* F \cup G$.

$E \succsim^* F$ holds if and only if for some $\{\alpha\} \succ \{\beta\}$, $\{\alpha\}_E\{\beta\} \succsim \{\alpha\}_F\{\beta\}$. By Axiom 7, this is equivalent to $\{\alpha\}_E\{\beta\}_G\{\beta\} \succsim \{\alpha\}_F\{\beta\}_G\{\beta\}$. By Axiom 2, this is equivalent to $\{\alpha\}_E\{\alpha\}_G\{\beta\} \succsim \{\alpha\}_F\{\alpha\}_G\{\beta\}$. By Axiom 7 again, this is equivalent to $\{\alpha\}_{E\cup G}\{\beta\} \succsim \{\alpha\}_{F\cup G}\{\beta\}$ and thus $E \cup G \succsim^* F \cup G$. $\square$

Lemma 13 (Quantitative Probability, (Villegas, 1964))**.** *If $\mathcal{E}$ has no atoms and $\succsim^*$ is a qualitative probability, then there exists a unique probability measure $\mu : \mathcal{E} \to [0, 1]$ that represents $\succsim^*$.*

Proof. By Villegas (1964), Theorem 3. □

Lemma 14 (Probability Weighted Representation). *$\succsim$ can be represented by:*

$$U(a) = \sum_{E \in \iota(a)} \mu(E) v\big(a(E), E\big) \tag{19}$$

with $v\big(\{\alpha\}, E\big) = v\big(\{\alpha\}, F\big)$ for all nonnull events E, F and all outcomes $\alpha \in \mathcal{X}$.

Proof. (sketch) We first show that $\mu(E) = 0$ if and only if E is null. Since $\mu(\emptyset) = 0$, it suffices to show $E \sim^* \emptyset$ if and only if E is null. E is null if and only if for all outcomes, $\alpha_E \beta \sim \beta$. By Axiom 11, $E \sim^* \emptyset$.

From the additive representation in the previous Lemma and that $\mu(E) = 0$ only if E is null, it follows that we can rewrite the representation into the desired form. It remains to show that $v(E, x) = v(F, x)$.

We claim that we can choose v such that $v(E, x) = v(F, x)$. If acts are restricted to singleton menus, by Axiom 7, we can show that the premises of Savage's theorem are satisfied. Hence, we have an expected utility $\sum_{E \in \iota(a)} \mu(E) w(a(E))$. Consider two information partitions P and Q that contain E. Fix the partition P first. Then both $\sum_{G \in P} \mu(G) v(G, a(G))$ and $\sum_{G \in P} \mu(G) w(a(G))$ are additive representations of the preference relation restricted to $\{a | \iota(a) = P \wedge \forall E \in P : |a(E)| = 1\}$, the set of all acts that consist of singleton menus and information partition P. Thus, $v(G, a(G)) = \sigma_P w(a(G)) + \tau_{P,G}$. In particular, $v(E, x) = \sigma_P w(x) + \tau_{P,E}$. Similarly, if we fix the partition Q, then $v(E, x) = \sigma_Q w(x) + \tau_{Q,E}$. Thus, we have

$$\sigma_P \Big(w(x) - w(y) \Big) = v(E, x) - v(E, y) = \sigma_Q \Big(w(x) - w(y) \Big).$$

Hence, $\sigma_P = \sigma_Q \equiv \sigma$ and $\tau_{P,E} = v(E, x) - \sigma w(x) = \tau_{Q,E} \equiv \tau_E$. Finally, let $\hat{v} = \frac{v - \tau_E}{\sigma}$. Then $\hat{v}(E, x) = w(x)$, which is independent of E and thus $\hat{v}(E, x) = \hat{v}(F, x)$. Since $\hat{v}$ is an affine transformation of v, it still represents the preference relation. □

Using an analogous proof to proposition 1, we obtain that each $v(F, m) = \max_{b \in m} \mu(F) v(F, \{b\})$. The instrumental knowledge property in turn guarantees that $\mu(F) v(F, \{b\}) = \sum_x \mu(F \cap b^{-1}(x)) v(F, \{x\})$. Since we have shown above that $v(F, \{x\})$ does not depend on events, it can be replaced by a suitably chosen utility function over outcomes $u(x)$.

G Monotone Sequences of Acts and Events

Lemma 15. *If $E^k \nearrow E$ or $E^k \searrow E$, then $E^k \rightsquigarrow E$.*

Proof. Suppose $E^k \nearrow E$. Since $E^k \subseteq E$ for all k, the second condition holds. Assume by contradiction that the first condition does not hold, i.e., there exists nonnull event $F \subsetneq E$ and for each N, there exists $k > N$ such that $F \cap E^k = \emptyset$. Since (E^k) is non-decreasing, this means $F \cap E^k = \emptyset$ for all k and thus $F \cap E = \emptyset$, a contradiction. Hence, the first condition holds.

Suppose $E^k \searrow E$. Since $E \subseteq E^k$ for all k, the first condition hold. Since (E^k) is non-increasing, we have if $F \subsetneq E^N$, then $F \subsetneq E^l$ for all $l < N$. Thus, if F satisfies the premise of the second condition, then $F \subsetneq E^k$ for all k. Hence, $F \cap E = \emptyset$ and the second condition holds. □

Lemma 16. *Let (E^k) be a sequence of events. Suppose for all ε, there exists N such that $\hat{\mu}(E^i \triangle E^j) < \varepsilon$ if $i, j > N$. Then (E^k) converges in subject belief to* $\limsup_{k\to\infty} E^k$.

Proof. Recall Definition 4. Suppose that the first condition does not hold. That is, there exists a nonnull $F \subsetneq \limsup_{k\to\infty} E^k$ and for each N, there exists $k > N$ such that $F \cap E^k = \emptyset$. Construct accordingly a subsequence (E^l) satisfying $F \cap E^l = \emptyset$ for all l. Since $F \subsetneq \limsup_{k\to\infty} E^k$, we can construct another subsequence $(E^{l'})$ satisfying $F \subseteq E^{l'}$ for all l'. If $E^i \in (E^l)$ and $E^j \in (E^{l'})$, then $\mu(E^i \triangle E^j) \geq \mu(F) > 0$, a contradiction.

To see the second condition holds, pick a nonnull F satisfying the requirement. Observe that $F \subseteq \bigcup_{k=N}^{\infty} E^k$ for all N. Hence, $F \subseteq \bigcap_{N=1}^{\infty} \bigcup_{k=N}^{\infty} E^k = \limsup_{k\to\infty} E^k$. □

Lemma 17. *$E^k \rightsquigarrow E$ if and only if E^k converge in measure $\hat{\mu}$ to E.*

Proof.

$$\lim_{k\to\infty} \hat{\mu}\big(\{s \in \mathcal{S} : |\mathbb{1}_{E^k}(s) - \mathbb{1}_E(s)| > \varepsilon\}\big) = 0.$$

$$\{s \in \mathcal{S} : |\mathbb{1}_{E^k}(s) - \mathbb{1}_E(s)| > \varepsilon\} = E^k \triangle E.$$

$$\lim_{k\to\infty} \hat{\mu}(E^k \triangle E) = 0.$$

($\Longrightarrow$) Suppose $\lim_{k\to\infty} \hat{\mu}(E^k \triangle E) = 0$. Pick nonnull $F \subsetneq E$. Assume by contradiction that (E1) does not hold. Then for each N, there exists $i > N$ such that $F \cap E^i = \emptyset$. Construct a subsequence (E^i) accordingly. Then $\hat{\mu}(E^i \triangle E) \geq \hat{\mu}(E \setminus E^i) \geq \hat{\mu}(F)$ for all i, a contradiction.

Assume by contradiction that (E2) does not hold. Then there exists a nonnull F and a subsequence (E^i) such that $F \subsetneq E^i$ for all i and $F \cap E = \emptyset$. This implies $\hat{\mu}(E^i \triangle E) \geq \hat{\mu}(E^i \setminus E) \geq \hat{\mu}(F)$ for all i, a contradiction.

($\Longleftarrow$) Suppose $E^k \rightsquigarrow E$. Assume by contradiction that $\hat{\mu}(E^k \setminus E)$ does not converge to 0. Then there exists ε and a subsequence (E^i) such that $\hat{\mu}(E^i \setminus E) > \varepsilon$ for all i. Consequently, for all i,

$$\hat{\mu}\left(\bigcup_{n \geq i} E^n \setminus E\right) \geq \hat{\mu}(E^i \setminus E) > \varepsilon.$$

Since the above inequality holds for all i, we have

$$\hat{\mu}(\limsup E^i \setminus E) > \varepsilon.$$

This contradicts to (E2) by picking $F = \limsup E^i \setminus E$.

Similarly, assume by contradcition that $\hat{\mu}(E \setminus E^k)$ does not converge to 0. Then there exists ε and a subsequence (E^j) such that $\hat{\mu}(E \setminus E^j) > \varepsilon$ for all j. Consequently, for all j,

$$\hat{\mu}\left(E \setminus \bigcap_{n \geq j} E^n\right) \geq \hat{\mu}(E \setminus E^j) > \varepsilon.$$

Since the above inequality holds for all j, we have

$$\hat{\mu}(E \setminus \liminf E^j) > \varepsilon.$$

This contradicts to (E1) by picking $F = E \setminus \liminf E^j$.

□

References

Blackwell, D. (1953). Equivalent Comparisons of Experiments. *The Annals of Mathematical Statistics*, *24*(2), 265–272. https://doi.org/10.1214/aoms/1177729032

Dekel, E., Lipman, B. L., & Rustichini, A. (2001). Representing preferences with a unique subjective state space. *Econometrica*, *69*(4), 891–934. Retrieved January 6, 2016, from http://onlinelibrary.wiley.com/doi/10.1111/1468-0262.00224/full

Dillenberger, D., Lleras, J. S., Sadowski, P., & Takeoka, N. (2014). A theory of subjective learning. *Journal of Economic Theory*, *153*, 287–312. https://doi.org/10.1016/j.jet.2014.07.003

Gilboa, I., & Lehrer, E. (1991). The value of information - An axiomatic approach. *Journal of Mathematical Economics*, *20*(5), 443–459. https://doi.org/10.1016/0304-4068(91)90002-B

Kreps, D. M. (1979). A Representation Theorem for "Preference for Flexibility". *Econometrica*, *47*(3), 565. https://doi.org/10.2307/1910406

Liang, Y. (2019). Information-dependent expected utility. http://ssrn.com/abstract-2842714

Morris, S. (1997). Risk, uncertainty and hidden information. *Theory and Decision*, *42*, 235–269.

Nehring, K. (1999). Preference for Flexibility in a Savage Framework. *Econometrica*, *67*, 121–146.

Rommeswinkel, H., Hsu, W.-T., & Chang, H.-C. (2020). Preference for Knowledge.

Savage, L. J. (1954). *The foundations of statistics*. Wiley.

van Zandt, T. (1996). Hidden information acquisition and static choice. *Theory and Decision*, *40*, 235–247.

Villegas, C. (1964). On Qualitative Probability $/sigma$-Algebras. *The Annals of Mathematical Statistics*, *35*(4), 1787–1796. https://doi.org/10.1214/aoms/1177700400

Wakker, P. P. (1989). *Additive representations of preferences: A new foundation of decision analysis* (Vol. 4). Springer Science & Business Media.